

## MATRIX OPTIMIZATION UNDER RANDOM EXTERNAL FIELDS

AMIR DEMBO\* AND OFER ZEITOUNI†

ABSTRACT. We consider the quadratic optimization problem

$$F_n^{W,\mathbf{h}} := \sup_{\mathbf{x} \in S^{n-1}} \left( \frac{1}{2} \mathbf{x}^T W \mathbf{x} + \mathbf{h}^T \mathbf{x} \right),$$

with  $W$  a (random) matrix and  $\mathbf{h}$  a random external field. We study the probabilities of large deviation of  $F_n^{W,\mathbf{h}}$  for  $\mathbf{h}$  a centered Gaussian vector with i.i.d. entries, both conditioned on  $W$  (a general Wigner matrix), and unconditioned when  $W$  is a GOE matrix. Our results validate (in a certain region) and correct (in another region), the prediction obtained by the mathematically non-rigorous replica method in Y. V. Fyodorov, P. Le Doussal, *J. Stat. phys.* **154** (2014).

## 1. INTRODUCTION

We consider in this paper the following quadratic optimization problem: given an  $n$ -by- $n$  symmetric matrix  $W$  and a vector  $\mathbf{h} \in \mathbb{R}^n$ , define, for  $\mathbf{x} \in S^{n-1} = \{\mathbf{x} : \sum_{i=1}^n x_i^2 = 1\}$  in the  $n$ -sphere, the quantity

$$E_n^{W,\mathbf{h}}(\mathbf{x}) := \frac{1}{2} \mathbf{x}^T W \mathbf{x} + \mathbf{h}^T \mathbf{x}, \quad (1.1)$$

and consider the optimization problem

$$F_n^{W,\mathbf{h}} := \sup_{\mathbf{x} \in S^{n-1}} \{E_n^{W,\mathbf{h}}(\mathbf{x})\}. \quad (1.2)$$

As discussed in [FLD13] to which we refer the reader for motivation and background, the quantity  $E_n^{W,\mathbf{h}}(\mathbf{x})$  has a natural interpretation as minus the energy associated with a configuration  $\mathbf{x}$  of  $n$  spin variables  $x_i$  in the presence of quadratic interaction  $W$  and an external field  $\mathbf{h}$ . In contrast to the situation when  $\mathbf{h} = 0$ , the function  $F_n^{W,\mathbf{h}}$  depends on the whole spectrum of  $W$  and not just on its top eigenvalue.

It is natural to consider both  $W$  and  $\mathbf{h}$  as random objects. Fixing  $\Gamma > 0$  constant, in [FLD13], the authors (among other things) use a mathematically non-rigorous replica method to study the large deviations of the random variable  $F_n^{W,\mathbf{h}}$  under the law  $\mathbb{P}_\Gamma^{G,n}$  where  $\mathbf{h}$  is a vector consisting of i.i.d. centered Gaussian variables of variance  $\Gamma/n$  and  $W$  is a matrix sampled from the Gaussian Orthogonal Ensemble (GOE). That is,  $W$  is a symmetric matrix whose entries  $\{W_{ij}\}$  on and above the diagonal are independent centered

---

*Date:* September 17, 2014.

*2010 Mathematics Subject Classification.* 60F10, 82D30.

*Key words and phrases.* Large deviations, replica method, random matrices, spin glass.

\* Stanford University. Research partially supported by NSF grant DMS-1106627.

† Weizmann Institute of Science and Courant Institute. Research partially supported by a grant from the Israel Science Foundation.

Gaussian variables of variance  $n^{-1}(1 + \delta_{ij})$ . In this setting, [FLD13] provides an argument for what we refer to below as an *annealed* Large Deviation Principle (LDP), in the following form (see [FLD13, formula (43)]).

**Prediction 1.1.** *Set  $m_c := \sqrt{1 + \Gamma/(1 + \Gamma)}$ . Then, for any  $m \in (m_c, \infty)$ ,*

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_\Gamma^{G,n} (|F_n^{W,\mathbf{h}} - m| < \delta) = -I_{FLD}^{G,A}(m; \Gamma), \quad (1.3)$$

where

$$I_{FLD}^{G,A}(m; \Gamma) = \frac{m}{1 + 2\Gamma} \left( -m\Gamma + (1 + \Gamma)\sqrt{m^2 - m_c^2} \right) - \log \left( \frac{\sqrt{1 + \Gamma}}{1 + 2\Gamma} \left( m + \sqrt{m^2 - m_c^2} \right) \right). \quad (1.4)$$

Note that no information is provided in Prediction 1.1 on what happens when  $m \leq m_c$ .

As mentioned above, the derivation in [FLD13] uses a non-rigorous replica trick and breaks down at  $m_c$ . Our interest in the problem was initiated by Y. Fyodorov, who asked whether Prediction 1.1 can be derived rigorously, and whether Prediction 1.1 can be extended to the regime  $m \leq m_c$ . This paper is devoted to answering these and related questions.

We find it advantageous and interesting to discuss first a *quenched* large deviations theorem, namely a large deviations statement when the sequence of matrices  $W = W_n$  is given. In this setup, the assumption that  $W$  is a GOE matrix (or, more generally, a Wigner matrix) plays no role. Under appropriate conditions summarized in Assumption 1.2, we derive in Theorem 1.3 a conditional (in  $\{W_n\}$ ) LDP for  $F_n^{W,\mathbf{h}}$  at speed  $n$  when the vector  $\mathbf{h}$  is either taken uniformly on  $\sqrt{\Gamma}S^{n-1}$  (with associated explicit Good Rate Function (GRF)  $I_q^H$ ), or when the entries of  $\mathbf{h}$  are i.i.d. centered Gaussians with variance  $\Gamma/n$  (with associated GRF  $I_q^G$ ). (See for example [DZ98, Sec. 1.2] for definitions of LDP and GRF). Theorem 1.3 then yields in a straightforward manner Corollary 1.7, which deals with general Wigner matrices (including, but not limited to, the GOE).

We then turn our attention to the case where  $W$  is sampled from the GOE. We derive the corresponding annealed (i.e. averaged on  $W$ ) LDP at speed  $n$  on the whole real line, see Corollary 1.9. The proof builds on our quenched LDP, together with the LDP for the top eigenvalue of Wigner matrices derived previously in [BDG01].

In the last subsection of the introduction, we simplify the general form of the quenched and annealed rate functions for the GOE. In particular, we show that Prediction 1.1 is only true for  $m > m_L := 1 + \frac{\Gamma}{2(1+\Gamma)}$ . Since  $m_L > m_c$ , this means that Prediction 1.1 does not hold in part of its domain (see Fig. 2 for a numerical example).

The annealed LDP for  $F_n^{W,\mathbf{h}}$  and our proof of it, are applicable more generally to any ensemble of random matrices having negligible fluctuations of their empirical spectral measures at our large deviations speed and scale, and for which the LDP at speed  $n$  of the maximal (or minimal), eigenvalue of  $W$  is available. In contrast with the *universality* of the rate functions  $I^H$  and  $I^G$  for the *quenched* large deviations of  $F_n^{W,\mathbf{h}}$ , the annealed rate functions  $I^{H,A}$  and  $I^{G,A}$  are specific to the GOE (as they depend on the exact form of the LDP for its maximal eigenvalue).

In the rest of the introduction we present the relevant notation and state our assumptions and main results.

**1.1. LDP for quadratic optimization problems.** Throughout we write  $\mathbf{x} = (x_1, \dots, x_n)$  for a vector in  $\mathbb{R}^n$  and  $\mathbf{x}^2 = (x_1^2, \dots, x_n^2)$ . The scalar product in  $\mathbb{R}^n$  is denoted  $\langle \cdot, \cdot \rangle$ , with  $\| \cdot \|$  for the Euclidean norm in  $\mathbb{R}^n$ . We further use  $M_+(J)$  for the space of all finite, Borel measures on  $J \subseteq \mathbb{R}$ , with  $M_1(J)$  denoting the sub-space of all probability measures on  $J$ , both equipped with the topology of weak convergence.

Let  $\mathbb{R}_{\geq}^n := \{(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}^n : \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n\}$  denote the collection of ordered  $n$ -tuple real numbers, with  $S^{n-1}$  denoting the usual Euclidean sphere of radius 1. For fixed  $\boldsymbol{\lambda} \in \mathbb{R}_{\geq}^n$  and constant  $\Gamma > 0$ , we are interested in large deviations for the (random) optimization problem

$$F_{n,\mathbf{h},\boldsymbol{\lambda}}^* = \sup_{\mathbf{x} \in S^{n-1}} \left( \frac{1}{2} \langle \boldsymbol{\lambda}, \mathbf{x}^2 \rangle + \langle \mathbf{h}, \mathbf{x} \rangle \right), \quad (1.5)$$

with respect to  $\mathbf{h} = (h_1, \dots, h_n) \in \sqrt{\Gamma} S^{n-1}$  a random vector drawn uniformly from the Haar measure on  $\sqrt{\Gamma} S^{n-1}$ , or alternatively when having  $\mathbf{h} = \mathbf{g}$  a centered multivariate normal random vector of covariance matrix  $\frac{\Gamma}{n} \mathbf{I}_n$ . Throughout we assume the following about the parameters of the optimization problem (1.5).

**Assumption 1.2.** *For  $n \rightarrow \infty$  we have that:*

- (A1).  $L_n^\lambda = \frac{1}{n} \sum_{j=1}^n \delta_{\lambda_j}$  converge weakly in  $M_1(\mathbb{R})$  to some  $q(\cdot)$  of compact support.
- (A2).  $\lambda_1(n) \rightarrow \lambda_+^* < \infty$  (necessarily,  $\lambda_+^* \geq q_+ := \max\{x \in \text{supp}(q)\}$ ).
- (A3).  $\lambda_n(n) \rightarrow \lambda_-^* > -\infty$  (necessarily,  $\lambda_-^* \leq q_- := \min\{x \in \text{supp}(q)\}$ ).

Our first result is then the following LDP.

**Theorem 1.3.** *Let Assumption 1.2 hold and fix  $\Gamma > 0$  non-random.*

(a). *For  $\mathbf{h}$  Haar distributed on  $\sqrt{\Gamma} S^{n-1}$ , the sequence  $\{F_{n,\mathbf{h},\boldsymbol{\lambda}}^*\}$  satisfies the LDP in  $\mathbb{R}$ , with speed  $n$  and GRF*

$$I_q^H(m; \lambda_\pm^*, \Gamma) = \inf \left\{ \frac{1}{2} H(q|\nu) : \nu \in M_1([\lambda_-^*, \lambda_+^*]), m = F(\lambda_+^*, \nu; \Gamma) \right\}, \quad (1.6)$$

where for given  $\xi \in \mathbb{R}$ ,

$$F(\xi, \nu; \Gamma) = \frac{1}{2} \inf_{\theta > \xi} \left[ \theta + \Gamma \int \frac{\nu(dx)}{\theta - x} \right], \quad (1.7)$$

for any  $\nu \in M_+((-\infty, \xi])$ , and

$$H(\mu|\nu) = \begin{cases} \int d\mu \log\left(\frac{d\mu}{d\nu}\right) + \nu(\mathbb{R}) - 1, & \mu \ll \nu, \\ \infty, & \text{otherwise.} \end{cases}$$

(b). *For  $\mathbf{g}$  centered multivariate normal of covariance  $\frac{\Gamma}{n} \mathbf{I}_n$ , the sequence  $\{F_{n,\mathbf{g},\boldsymbol{\lambda}}^*\}$  satisfies the LDP with speed  $n$  and the GRF*

$$\begin{aligned} I_q^G(m; \lambda_\pm^*, \Gamma) &= \inf \left\{ \frac{1}{2} H(q|\nu) : \nu \in M_+([\lambda_-^*, \lambda_+^*]), m = F(\lambda_+^*, \nu; \Gamma) \right\} \\ &= \inf_{y \geq 0} \{ I_q^H(m; \lambda_\pm^*, \Gamma y) + J_1(y) \}, \end{aligned} \quad (1.8)$$

where

$$J_1(y) = \frac{1}{2}(y - 1 - \log y), \quad \forall y \in \mathbb{R}_+. \quad (1.9)$$

**Remark 1.4.** Clearly,  $I_q^G(\cdot; \lambda_\pm^*, \Gamma) \leq I_q^H(\cdot; \lambda_\pm^*, \Gamma)$  and both of these GRFs are zero if and only if  $\nu = q$  and  $y = 1$ . That is, when  $m = \bar{m}$ , where

$$\bar{m} := F(\lambda_+^*, q; \Gamma). \quad (1.10)$$

Further, from (1.7) we see that  $F(\lambda_+^*, \nu; \Gamma) \in [m_-^*, m_+^*]$  for  $\nu \in M_1([\lambda_-^*, \lambda_+^*])$  and

$$m_-^* := F(\lambda_+^*, \delta_{\lambda_-^*}; \Gamma) = \frac{1}{2}(\theta_-^* + \frac{\Gamma}{\theta_-^* - \lambda_-^*}), \quad \text{where } \theta_-^* = \lambda_+^* \vee (\lambda_-^* + \sqrt{\Gamma}) \quad (1.11)$$

$$m_+^* := F(\lambda_+^*, \delta_{\lambda_+^*}; \Gamma) = \frac{1}{2}\lambda_+^* + \sqrt{\Gamma}. \quad (1.12)$$

Hence,  $I_q^H(\cdot; \lambda_\pm^*, \Gamma) = \infty$  outside the compact interval  $[m_-^*, m_+^*]$  (which is strictly above  $\frac{1}{2}\lambda_+^*$ ), whereas  $I_q^G(m; \lambda_\pm^*, \Gamma) = \infty$  for  $m \leq \frac{1}{2}\lambda_+^*$ .

We next detail a few regularity properties of the rate functions of Theorem 1.3.

**Proposition 1.5.** The GRF  $I_q^H(\cdot; \lambda_\pm^*, \Gamma)$  is continuous on  $(m_-^*, m_+^*)$ , non-increasing on  $(m_-^*, \bar{m}]$  and convex strictly increasing on  $[\bar{m}, m_+^*)$ , whereas the GRF  $I_q^G(\cdot; \lambda_\pm^*, \Gamma)$  is continuous on  $(\frac{1}{2}\lambda_+^*, \infty)$ , non-increasing on  $(\frac{1}{2}\lambda_+^*, \bar{m}]$  and convex strictly increasing on  $[\bar{m}, \infty)$ .

**Remark 1.6.** See also Proposition 1.10 for more explicit expressions for the rate functions  $I_q^H$  and  $I_q^G$ . In particular, it is shown there that  $I_q^G(m; \lambda_\pm^*, \Gamma)$  is independent of  $\lambda_-^*$ , as is  $I_q^H(m; \lambda_\pm^*, \Gamma)$  when  $m > \bar{m}$ , whereas  $I_q^H(m; \lambda_\pm^*, \Gamma)$  is independent of  $\lambda_+^*$  when  $m < \bar{m}$ .

**1.2. LDP for random quadratic forms - Wigner matrices versus the GOE.** The general LDP of Theorem 1.3 yields LDPs for quadratic optimization problems involving random matrices. Fixing  $\lambda_\pm^* \in \mathbb{R}$  and  $q \in M_1([\lambda_-^*, \lambda_+^*])$ , let  $\mathcal{W}_{\lambda_\pm^*, q}$  denote the collection of all sequences of (random or deterministic) symmetric  $n$ -dimensional  $\mathbb{R}$ -valued matrices, whose ordered eigenvalue vectors  $\lambda$  satisfy Assumption 1.2 for  $q$  and  $\lambda_\pm^*$ . The following LDP for  $\{F_n^{W, \mathbf{h}}\}$  is a direct consequence of Theorem 1.3. Here and in the sequel, for a sequence  $\{W_n, \mathbf{h}_n\}$  we write  $F_n^{W, \mathbf{h}}$  as shorthand for  $F_n^{W_n, \mathbf{h}_n}$ .

**Corollary 1.7** (Quenched LDP). Fix a deterministic constant  $\Gamma > 0$  and a sequence  $\{W_n\} \in \mathcal{W}_{\lambda_\pm^*, q}$ . For  $\tilde{\mathbf{h}}$  independent of  $W_n$ , denote by  $\mathbb{P}_\Gamma^{W, H, n}$  the law of  $F_n^{W, \tilde{\mathbf{h}}}$  conditioned on  $W_n$  if  $\tilde{\mathbf{h}}$  is Haar distributed on  $\sqrt{\Gamma}S^{n-1}$ , and by  $\mathbb{P}_\Gamma^{W, G, n}$  if  $\tilde{\mathbf{h}}$  is a centered multivariate normal  $\tilde{\mathbf{g}}$ , of covariance  $\frac{\Gamma}{n}\mathbf{I}_n$ .

(a). The sequence  $\{\mathbb{P}_\Gamma^{W, H, n}\}_{n \geq 1}$  satisfies the LDP on  $\mathbb{R}$  with speed  $n$  and the GRF  $I_q^H(m; \lambda_\pm^*, \Gamma)$  given by (1.6) (or alternatively, (1.18)).

(b). The sequence  $\{\mathbb{P}_\Gamma^{W, G, n}\}_{n \geq 1}$  satisfies the corresponding LDP with GRF  $I_q^G(m; \lambda_\pm^*, \Gamma)$  given by (1.8) (or alternatively, (1.22)).

**Remark 1.8.** Recall that a symmetric random matrix  $W_n$  is a Wigner matrix if it has centered independent entries on and above the diagonal, with the entries above the diagonal being i.i.d. of variance 1 and bounded fourth moment, while the on-diagonal entries are i.i.d. with uniformly bounded second moment. Such matrices are a.s. in  $\mathcal{W}_{\pm 2, \sigma}$ , with  $\sigma$

the semi-circle law having the support  $[-2, 2]$  and density  $f_\sigma(x) = (2\pi)^{-1}\sqrt{4-x^2}1_{|x|\leq 2}$ , see [AGZ10, Theorem 2.1.21] and [BY88]. Hence, all the conclusions of Corollary 1.7 hold for such matrices.

We turn to the LDP averaged over the choice of  $W_n$  from the GOE. Let  $\mathbb{P}_\Gamma^{H,n} = \mathbb{E}_{\text{GOE}} \mathbb{P}_\Gamma^{W,H,n}$  and  $\mathbb{P}_\Gamma^{G,n} = \mathbb{E}_{\text{GOE}} \mathbb{P}_\Gamma^{W,G,n}$ . Note that under either  $\mathbb{P}_\Gamma^{H,n}$  or  $\mathbb{P}_\Gamma^{G,n}$ , the matrix  $W$  is sampled according to the GOE and is independent of the random vector  $\mathbf{h}$ .

**Corollary 1.9** (Annealed LDP).

(a). The sequence  $\{\mathbb{P}_\Gamma^{H,n}\}_{n\geq 1}$  satisfies the LDP with speed  $n$  and GRF

$$I^{H,A}(m; \Gamma) = \inf_{\psi_-^* \leq -2, \psi_+^* \geq 2} \left\{ I_\sigma^H(m; \psi_\pm^*, \Gamma) + I_e(\psi_+^*) + I_e(-\psi_-^*) \right\}, \quad (1.13)$$

for  $I^H(\cdot)$  of (1.6) (or alternatively (1.18)), and

$$I_e(\psi) = \begin{cases} \int_2^\psi \sqrt{(u/2)^2 - 1} du, & \psi \geq 2, \\ \infty, & \text{otherwise.} \end{cases} \quad (1.14)$$

(b). The sequence  $\{\mathbb{P}_\Gamma^{G,n}\}_{n\geq 1}$  satisfies the LDP with speed  $n$  and GRF

$$I^{G,A}(m; \Gamma) = \inf_{\psi^* \geq 2} \left\{ I_\sigma^G(m; \pm\psi^*, \Gamma) + I_e(\psi^*) \right\}, \quad (1.15)$$

for  $I^G(\cdot)$  of (1.8) (or alternatively (1.22)).

**1.3. Explicit rate functions.** We shall derive explicit expressions for the various rate functions introduced in the article, starting with the GRF-s of Theorem 1.3, for general  $q \in M_1([\lambda_-^*, \lambda_+^*])$ . To state the result, we require the logarithmic potential and Stieltjes transform of  $q(\cdot)$ , denoted by

$$\mathbf{L}(\xi) = \int \log|\xi - x|q(dx), \quad \forall \xi \notin (\lambda_-^*, \lambda_+^*), \quad (1.16)$$

$$\mathbf{G}(\xi) = \mathbf{L}'(\xi) = \int (\xi - x)^{-1}q(dx), \quad \forall \xi \notin (\lambda_-^*, \lambda_+^*). \quad (1.17)$$

**Proposition 1.10** (Quenched rate functions).

(a) In case  $q_\pm = \lambda_\pm^*$ , the GRF for part (a) of Theorem 1.3 is

$$I_q^H(m; \lambda_\pm^*, \Gamma) = \frac{1}{2} [\log|B| + \mathbf{L}(\psi) - \mathbf{L}(\theta)]. \quad (1.18)$$

Here  $t \geq 0$ ,  $\theta \geq \lambda_+^*$ ,  $B$  and  $\psi$  are such that

$$B - 1 = (\theta - \psi)\mathbf{G}(\psi) + Bt \quad (1.19)$$

$$B(2m - \theta) = \Gamma\mathbf{G}(\psi) + \frac{B\Gamma t}{\theta - \psi^*} \quad \text{with } t = 0 \text{ whenever } \psi \neq \psi^*, \quad (1.20)$$

$$B(\psi - \theta) \geq \Gamma\mathbf{G}(\theta) - \Gamma\mathbf{G}(\psi) - \frac{B\Gamma t}{\theta - \psi^*} \quad \text{with equality whenever } \theta > \lambda_+^*, \quad (1.21)$$

with  $m > \bar{m}$  requiring  $B > 0$  and  $\theta > \psi \geq \psi^* = \lambda_+^*$ , while for  $m < \bar{m}$  we consider either  $B > 0$  and  $\psi > \theta$ , or  $B < 0$  and  $\psi \leq \psi^* = \lambda_-^*$ .

(b) For any  $q \in M_1([\lambda_-^*, \lambda_+^*])$ , the GRF for part (b) of Theorem 1.3 is

$$I_q^G(m; \lambda_\pm^*, \Gamma) = \frac{1}{2}[(\theta - \psi)\mathbf{G}(\psi) + t + \mathbf{L}(\psi) - \mathbf{L}(\theta)], \quad (1.22)$$

with  $B = 1$  and  $\theta, \psi \geq \psi^* = \lambda_+^*$ ,  $t \geq 0$ , are determined by (1.20) and (1.21).

Building on Proposition 1.10, due to the simple form of  $\mathbf{G}(\cdot)$  for the semi-circle, one can explicitly solve the variational problems in the definition of the rate functions

$$I^{G,A} \leq I^{H,A} \wedge I_\sigma^G \leq I^{H,A} \vee I_\sigma^G \leq I_\sigma^H. \quad (1.23)$$

To state the result, we introduce  $\alpha, \beta \geq 1$  such that  $\theta = \alpha + \alpha^{-1}$  and  $\psi = \beta + \beta^{-1}$ , the functions

$$\mathcal{I}(\alpha, \beta) = J_1\left(\frac{\alpha}{\beta}\right) - \frac{1}{4}(\alpha^{-1} - \beta^{-1})^2, \quad \mathfrak{T}(\alpha) = \Gamma^{-1}(\alpha + \frac{1}{\alpha} - 2)[2m - \alpha - \frac{1}{\alpha} - \Gamma], \quad (1.24)$$

and the constants  $1 < m_c < m_L < \bar{m} < m_U$ , given by

$$m_c = \sqrt{\frac{1+2\Gamma}{1+\Gamma}}, \quad m_L = 1 + \frac{\Gamma}{2(1+\Gamma)}, \quad \bar{m} = \sqrt{1+\Gamma}, \quad m_U = 1 + \frac{\Gamma(1+2\Gamma)}{2(1+\Gamma)}. \quad (1.25)$$

**Proposition 1.11** (Rate functions with semi-circle).

(a). The quenched GRF  $I_\sigma^G(\cdot; \pm 2, \Gamma)$  is given by the formula

$$I_\sigma^G(m; \pm 2, \Gamma) = \mathcal{I}(\alpha_q, \beta_q) + \frac{1}{2}\mathfrak{T}(\alpha_q)\mathbf{1}_{\{m \geq m_U\}}, \quad (1.26)$$

where  $\alpha_q \leq \beta_q$  iff  $m \leq \bar{m}$  are given by

$$(\alpha_q, \beta_q) = \begin{cases} (1, \frac{\Gamma}{2(m-1)}), & m \in (1, m_L], \end{cases} \quad (1.27)$$

$$(\alpha_q, \beta_q) = \begin{cases} (m_c^{-2}[m + \sqrt{m^2 - m_c^2}], (1+\Gamma)[m - \sqrt{m^2 - m_c^2}]), & m \in (m_L, m_U), \end{cases} \quad (1.28)$$

$$(\alpha_q, \beta_q) = \begin{cases} (\frac{1}{2}[(m+1) + \sqrt{(m+1)^2 - 4 - 2\Gamma}], 1), & m \in [m_U, \infty). \end{cases} \quad (1.29)$$

(b). The strictly convex annealed GRF  $I^{G,A}(m; \Gamma)$  equals the quenched GRF  $I_\sigma^G(m; \pm 2, \Gamma)$  from (1.26) for  $m \in [1, m_U]$ , whereas for  $m > m_U$ ,

$$I^{G,A}(m; \Gamma) = \mathcal{I}(\alpha_a, \beta_a^{-1}) \quad (1.30)$$

for  $(\alpha_a, \beta_a^{-1})$  as in (1.28), i.e.

$$(\alpha_a, \beta_a^{-1}) = (m_c^{-2}[m + \sqrt{m^2 - m_c^2}], (1+\Gamma)[m - \sqrt{m^2 - m_c^2}]). \quad (1.31)$$

See Figure 1 for a plot of the quenched and annealed rate functions  $I_\sigma^G(m; \pm 2, 1)$  and  $I^{G,A}(m; 1)$ .

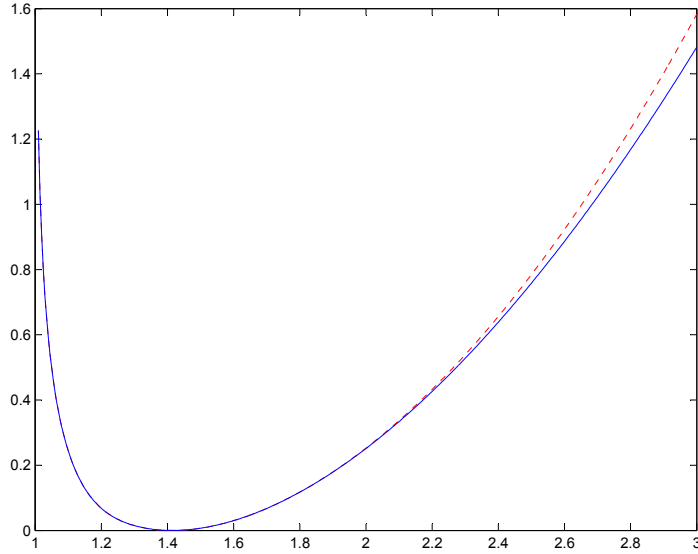


FIGURE 1. The rate functions  $I^{G,A}(m; 1) < I_{\sigma}^G(m; \pm 2, 1)$  (here  $m_L = 1.25$ ,  $m_U = 1.75$ )

**Remark 1.12.** The annealed GRF  $I^{G,A}(m; \Gamma)$  could also be written as

$$I^{G,A}(m; \Gamma) = \mathcal{I}(\alpha, \beta)$$

where  $(\alpha, \beta) = (\alpha_q, \beta_q)$  if  $m \leq m_U$ , and  $(\alpha, \beta)$  given by the r.h.s. of (1.31), if  $m > m_U$ . Note that while  $m \mapsto I^{G,A}(m; \Gamma)$  is smooth except for the jump discontinuity of its third derivative at  $m = m_L$ , the function  $m \mapsto I_{\sigma}^G(m; \pm 2, \Gamma)$  is also non-smooth at  $m = m_U$ .

**Remark 1.13.** It is worthwhile to comment on the relation between the rate function  $I^{G,A}(m; \Gamma)$  of Proposition 1.11 and Prediction 1.1 from [FLD13]: a tedious, but straight forward algebraic manipulation shows that  $I^{G,A}(m; \Gamma) = I_{FLD}^{G,A}(m; \Gamma)$  for  $m \geq m_L$ . However, both a numerical evaluation, see Figure 2, and analytic evaluation of the limit  $m \searrow m_c$  as well as comparison of the first three derivatives at  $m = m_L$ , show that in general  $I^{G,A}(m; \Gamma) \neq I_{FLD}^{G,A}(m; \Gamma)$  in the interval  $m \in (m_c, m_L)$ .

We conclude the introduction with comments on the rate functions  $I_{\sigma}^G$  and  $I^{G,A}$ . The general form of the rate function can be understood by considering the following heuristics. Three main objects enter the (diagonalized) optimization problem (1.5):

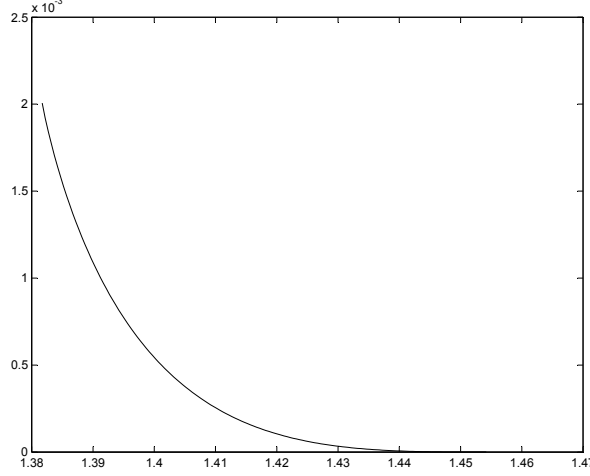


FIGURE 2.  $I_{FLD}^{G,A}(m; 10) - I^{G,A}(m; 10)$  for  $m \in (m_c, m_L)$ ; here  $m_c = 1.38[2]$ ,  $m_L = 1.45[4]$

- (1) The total mass  $\sum_{i=1}^n h_i^2$ , which we take to roughly equal  $\Gamma y$ .
- (2) The measure of total mass  $y > 0$ , which controls the distribution of the  $h_i^2$  as weights on the eigenvalues  $\lambda_i$ , denoted  $\nu = \Gamma^{-1} \sum_{i=1}^n h_i^2 \delta_{\lambda_i}$ .
- (3) The optimal profile of  $x_i$ -s for a given  $\nu$ , which turns out to be determined by a Lagrange multiplier  $\theta \geq \lambda_+^*$  (specifically,  $x_i = h_i/(\theta - \lambda_i)$ , as shown in Lemma 2.2).

The minimization of the probabilistic cost of producing such  $\nu$  while constraining the value  $F_{n,\mathbf{h},\boldsymbol{\lambda}}^* \approx m$  yields for  $n \rightarrow \infty$  the optimal

$$\nu^*(d\lambda) = \frac{\theta - \lambda}{\psi - \lambda} q(d\lambda) + t \delta_{\lambda_+^*} \quad (1.32)$$

in terms of another Lagrange multiplier, denoted  $\psi \geq \lambda_+^*$  (see proof of part (b) of Proposition 1.10 for the derivation of  $\nu^*$ ). We note in passing that  $t = t(m)$  represents the total mass projected by  $\mathbf{h}$  on the eigenspace of  $o(n)$  top eigenvalues of  $W$ , if constrained to  $F_n^{W,\mathbf{h}} \approx m$ , and is non-zero only when  $\psi$  is at the edge of the vector  $\boldsymbol{\lambda}$ . Now the three regimes of the quenched rate function in Proposition 1.11, where  $\lambda_+^* = 2$ , correspond to the following cases:

$$\begin{aligned} m \in (1, m_L] &\iff \psi > \theta = \lambda_+^*, & t = 0 \\ m \in (m_L, m_U) &\iff \theta > \lambda_+^*, \psi > \lambda_+^*, & t = 0 \\ m \in [m_U, \infty) &\iff \theta > \psi = \lambda_+^*, & t > 0. \end{aligned}$$

At the typical value  $\bar{m}$  which lies in  $(m_L, m_U)$ , one switches from having  $\psi > \theta$  (lower tail large deviations, with  $y = y(m) < 1$ ), to  $\theta > \psi$  (upper tail large deviations, with



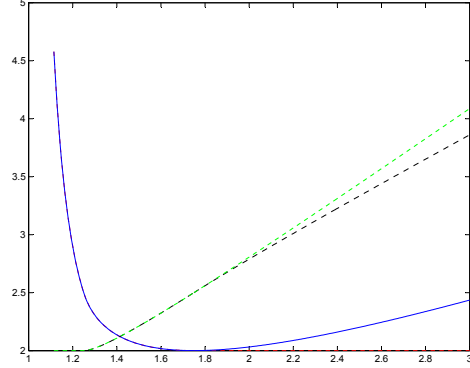


FIGURE 3. The parameters  $\psi(m)$  (quenched, dashed red; annealed, solid blue) and  $\theta(m)$  (quenched, dashed black; annealed, dotted green), for  $\Gamma = 1$ .

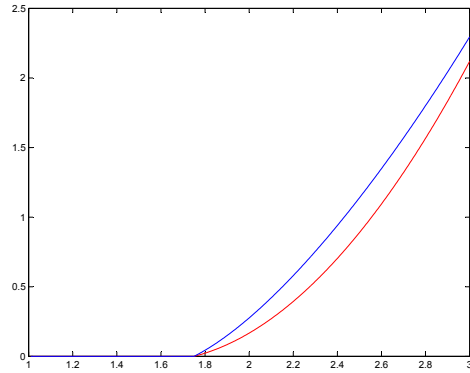


FIGURE 4. The parameter  $t(m)$  (quenched, dashed black; annealed, dotted green), at  $\Gamma = 1$ .

$y = y(m) > 1$ ). In the annealed case described in Proposition 1.11, the regime  $m \in [m_U, \infty)$  is different because while saturating the constraint on  $\psi$  at the value of the top eigenvalue, the optimal solution is now able to shift the top eigenvalue from  $\lambda_+^*$  to  $\psi = \psi(m) > \lambda_+^*$ . See Figures 3-4 for a plot of the parameters of (1.32) in the quenched and annealed cases (at  $\Gamma = 1$ , with Fig. 3 depicting  $m \mapsto \psi(m)$ ,  $m \mapsto \theta(m)$ , and Fig. 4 for  $m \mapsto t(m)$ ).

**Remark 1.14.** *The parameter  $\Gamma$  determines the relevant importance of the quadratic and linear parts of the optimization problem (1.5). In one extreme of  $\Gamma \gg 1$ , the typical value*

is  $\bar{m} = \sqrt{\Gamma}(1 + o(1))$  with (1.5) dominated by its linear (Gaussian) part. In this case, except for the extreme tails (i.e.  $m \leq m_L \approx 3/2$  or  $m \geq m_U \approx \Gamma$ ), both  $I^{G,A}(m; \Gamma)$  and  $I_\sigma^G(m; \pm 2, \Gamma)$  approximately match the  $\chi$ -square rate  $J_1(y)$  for  $y = m^2/\Gamma$  (with the contribution of eigenvalues buried in the correction terms). In contrast, for  $\Gamma \downarrow 0$  the quadratic part dominates. Further, with  $m_U(\Gamma) \downarrow 1$ , it's all about the top of the spectrum of  $W_n$ . Here  $I^{G,A}(m; \Gamma)$  approximately matches the GOE rate function  $I_e(2m)$  (and since  $\mathbf{h}$  does not matter much, one uses  $t(m)$  small and  $\theta(m) \approx \psi(m) \approx 2m$  to get there), whereas  $I_\sigma^G(m; \pm 2, \Gamma) \approx t/2 = (m-1)^2/(2\Gamma)$  is the cost of making up the  $m-1$  discrepancy in the value of  $F_n^{W, \mathbf{h}}$  by having  $\mathbf{h}$  of that magnitude, aligned to the top eigenvector of  $W_n$ .

## 2. PROOFS

### 2.1. Rate functions: regularity properties.

*Proof of Proposition 1.5.* Fixing  $\lambda_\pm^*$  and  $\Gamma$ , let  $\mathcal{M} = M_+([\lambda_-^*, \lambda_+^*])$ ,  $\mathcal{M}_1 := M_1([\lambda_-^*, \lambda_+^*])$ , and  $f(\theta, \nu) := \frac{1}{2}(\theta + \Gamma \int (\theta - x)^{-1} \nu(dx))$  for  $\theta \geq \lambda_+^*$  and  $\nu \in \mathcal{M}$ . Consequently,

$$F(\nu) := F(\lambda_+^*, \nu; \Gamma) = \inf_{\theta > \lambda_+^*} \{f(\theta, \nu)\}, \quad (2.1)$$

which must be in  $[m_-^*, m_+^*]$  when  $\nu \in \mathcal{M}_1$ , since then  $f(\theta, \nu) \in [f(\theta, \delta_{\lambda_-^*}), f(\theta, \delta_{\lambda_+^*})]$ . As  $f(\theta, \nu) \rightarrow \infty$  when  $\theta \rightarrow \infty$ , by monotone convergence, for any  $\nu \in \mathcal{M}$  there exists some  $\theta_\nu \geq \lambda_+^*$  such that  $F(\nu) = f(\theta_\nu, \nu)$ .

**I. Continuity of  $F(\cdot)$ .** The map  $\nu \mapsto f(\theta, \nu)$  is continuous on  $\mathcal{M}$  for each fixed  $\theta > \lambda_+^*$ , so by (2.1) the infimum  $F(\nu)$  of these maps is upper semicontinuous (u.s.c.) on  $\mathcal{M}$ . To show that  $F(\nu)$  is lower semi continuous (l.s.c.), fix a convergent sequence  $\nu_n \rightarrow \nu$  in  $\mathcal{M}$  and let  $\theta_n := \theta_{\nu_n}$ . Passing to a sub-sequence, we may and will assume w.l.o.g. that  $\theta_n \rightarrow \theta^*$  for some  $\theta^*$  finite. If  $\theta^* > \lambda_+^*$  then the continuous functions  $x \rightarrow (\theta_n - x)^{-1}$  on the compact  $[\lambda_-^*, \lambda_+^*]$  converge uniformly to the continuous function  $x \rightarrow (\theta^* - x)^{-1}$ , from which we deduce that as  $n \rightarrow \infty$ ,

$$F(\nu_n) = f(\theta_n, \nu_n) \rightarrow f(\theta^*, \nu) \geq F(\nu)$$

(using (2.1) to get the inequality). Alternatively, if  $\theta_n \rightarrow \lambda_+^*$  then  $\theta_n \leq \lambda_+^* + 2\delta$  for any fixed  $\delta > 0$  and all  $n$  large enough, in which case by monotonicity of  $\theta \mapsto (\theta - x)^{-1}$  and the preceding argument, we have that as  $n \rightarrow \infty$ ,

$$F(\nu_n) + \delta = f(\theta_n, \nu_n) + \delta \geq f(\lambda_+^* + 2\delta, \nu_n) \rightarrow f(\lambda_+^* + 2\delta, \nu) \geq F(\nu).$$

Considering  $\delta \rightarrow 0$  yields the stated l.s.c., hence continuity, of  $F(\cdot)$ .

**II. The finiteness of  $I_q^H$  and  $I_q^G$ .** Setting now  $q_t^\pm := t\delta_{\lambda_\pm^*} + (1-t)q$ , both  $F(q_t^+) : [0, 1] \mapsto [\bar{m}, m_+^*]$  and  $F(q_t^-) : [0, 1] \mapsto [m_-^*, \bar{m}]$  are continuous in  $t$ , so by the mean-value theorem, for any  $m \in (m_-^*, m_+^*)$  there exists  $t = t(m) \in [0, 1)$  such that  $F(q_t^\pm) = m$ . Since  $H(q|q_t^\pm)$  are finite, so is  $I_q^H(m) := I_q^H(m; \lambda_\pm^*, \Gamma)$ . As for its boundary points, note that  $I_q^H(m_\pm^*) = \infty$  unless  $q = \delta_{\lambda_\pm^*}$ , in which case  $\bar{m} = m_\pm^*$ . Similarly,  $F(yq) : \mathbb{R}_+ \mapsto [\frac{1}{2}\lambda_+^*, \infty)$  is continuous in  $y$ , hence for any  $m > \frac{1}{2}\lambda^*$  there exists  $y = y(m) > 0$  such that  $F(yq) = m$ . With  $H(q|yq) = 2J_1(y)$  finite at any such  $y(m)$ , we deduce from the first identity in (1.8) that  $I_q^G(m) := I_q^G(m; \lambda_\pm^*, \Gamma)$  is finite for all  $m > \frac{1}{2}\lambda_+^*$ .

**III. Monotonicity and convexity of  $I_q^H$  and  $I_q^G$ .** Both  $I_q^H(m)$  and  $I_q^G(m)$  are non-decreasing for  $m \geq \bar{m}$  in their respective domains. Indeed, as seen in step II, there is no need to consider the boundary points. So, setting either  $I_q = I_q^H$  or  $I_q = I_q^G$  and fixing  $m' \geq m$  in the interior of the relevant interval, for any  $\epsilon > 0$  there exists  $\nu \in \mathcal{M}$  such that  $F(\nu) = m'$  and  $\frac{1}{2}H(q|\nu) \leq I_q(m') + \epsilon$  (with  $\nu \in \mathcal{M}_1$  in case  $I_q = I_q^H$ ). By the continuity of  $F(t\nu + (1-t)q) : [0, 1] \mapsto [\bar{m}, m']$  we have that  $F(\bar{\nu}) = m$  for  $\bar{\nu} = s\nu + (1-s)q$  and some  $s \in [0, 1]$ . Hence, by the convexity of  $\nu \mapsto H(q|\nu)$ ,

$$I_q(m) \leq \frac{1}{2}H(q|\bar{\nu}) \leq \frac{1}{2}(sH(q|\nu) + (1-s)H(q|q)) \leq I_q(m') + \epsilon.$$

The claimed monotonicity, namely  $I_q(m) = \inf_{m' \geq m} I_q(m')$ , follows upon considering  $\epsilon \downarrow 0$  (and by the same reasoning we also get that  $I_q(m)$  is non-increasing for  $m \leq \bar{m}$ ). This monotonicity further results with the convexity of  $I_q(m)$  for  $m \geq \bar{m}$ . Indeed, for such values of  $m$  we have that

$$I_q^G(m) = \inf \left\{ \frac{1}{2}H(q|\nu) : \nu \in \mathcal{M}, F(\nu) \geq m \right\},$$

with the analogous formula for  $I_q^H(m)$ , just requiring then to also have  $\nu \in \mathcal{M}_1$ . Now, by the concavity of  $F(\cdot)$ , if  $H(q|\nu_i) \leq 2I_q(m_i) + \epsilon$  and  $F(\nu_i) \geq m_i$ ,  $i = 1, 2$ , then for any  $s \in [0, 1]$ , both  $F(s\nu_1 + (1-s)\nu_2) \geq sm_1 + (1-s)m_2$  and  $H(q|s\nu_1 + (1-s)\nu_2) \leq 2(sI_q(m_1) + (1-s)I_q(m_2)) + \epsilon$ , implying that  $I_q(sm_1 + (1-s)m_2) \leq sI_q(m_1) + (1-s)I_q(m_2)$  (upon taking  $\epsilon \downarrow 0$ ). Finally, since  $I_q(m)$  is zero only at  $m = \bar{m}$  and convex at all  $m \geq \bar{m}$ , it must be strictly increasing at any  $m \geq \bar{m}$  in its domain.

**IV. The continuity of  $I_q^H$  and  $I_q^G$ .** Clearly the strictly increasing  $I_q(m) \rightarrow \infty$  when  $m \rightarrow \infty$ . Thus, the non-negative  $I_q(\cdot)$  is a GRF provided it is l.s.c. throughout  $\mathbb{R}_+$ , and to show such l.s.c. it suffices to consider  $m_n \rightarrow m$  for which  $\alpha := \liminf_{n \rightarrow \infty} I_q(m_n)$  is finite. Recall that  $J_1(y) \rightarrow \infty$  as  $y \rightarrow \infty$ , so in view of (1.8) we may always restrict our attention to a compact subset  $\mathcal{M}_{[0, y]} = \{\nu \in \mathcal{M} : \nu(\mathbb{R}) \leq y\}$  of  $\mathcal{M}$  for some  $y = y(\alpha) \geq 1$  large enough. Then, by the continuity of  $F(\cdot)$  on the compact  $\mathcal{M}_{[0, y]}$  we can pass to a sub-sequence  $\{n_k\}$  for which there exist  $\nu_{n_k} \rightarrow \nu \in \mathcal{M}_{[0, y]}$  with  $F(\nu_{n_k}) = m_{n_k}$  and  $\frac{1}{2}H(q|\nu_{n_k}) \rightarrow \alpha$ . Since  $m_{n_k} = F(\nu_{n_k}) \rightarrow F(\nu)$  it follows that  $F(\nu) = m$  and consequently by the l.s.c. of  $\nu \mapsto H(q|\nu)$ ,

$$\alpha = \frac{1}{2} \lim_{k \rightarrow \infty} H(q|\nu_{n_k}) \geq \frac{1}{2}H(q|\nu) \geq I_q(m),$$

as claimed. The continuity of  $I_q(\cdot)$  at any  $m > \bar{m}$  in the interior of its domain, follows from the convexity of  $I_q(\cdot)$ . With  $m \mapsto I_q(m)$  non-increasing at any  $m \leq \bar{m}$ , it suffices to fix  $\epsilon > 0$  and  $m \leq \bar{m}$  with  $I_q(m) < \infty$  and show the existence of  $m_n \nearrow m$  such that  $\liminf_n I_q(m_n) \leq I_q(m) + \epsilon$ . To this end, there exists  $\nu \in \mathcal{M}$ ,  $\nu \neq 0$ , such that  $F(\nu) = m$  and  $H(q|\nu) \leq 2I_q(m) + \epsilon$  (further having  $\nu \in \mathcal{M}_1$ ,  $\nu \neq \delta_{\lambda_-^*}$  in case  $I_q = I_q^H$ ). Then, setting  $\nu_t = (1-t)\nu + t\delta_{\lambda_-^*} \mathbf{1}_{\{I_q = I_q^H\}}$  we have that  $t \mapsto F(\nu_t)$  is continuous, with  $H(q|\nu_t) \leq H(q|\nu) - \log(1-t)$  and  $F(\nu_t) < F(\nu_0) = m$  for all  $t \in (0, 1]$ . Hence, fixing any  $t_n \downarrow 0$  results with  $m_n = F(\nu_{t_n}) \nearrow m$ , such that

$$\liminf_{n \rightarrow \infty} I_q(m_n) \leq \frac{1}{2} \liminf_{n \rightarrow \infty} H(q|\nu_{t_n}) \leq \frac{1}{2}H(q|\nu) \leq I_q(m) + \epsilon$$

as needed for completing the proof. ■

**2.2. A finite dimensional optimization problem.** For  $K > 1$  integer,  $\mathbf{h} \in \sqrt{\Gamma}S^{K-1}$ ,  $\boldsymbol{\lambda} \in \mathbb{R}_{\geq}^K$  and  $\mathbf{x} \in \mathbb{R}^K$ , define

$$F_{K,\mathbf{h},\boldsymbol{\lambda}}(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^K \lambda_i x_i^2 + \sum_{i=1}^K h_i x_i.$$

Our next proposition provides an alternative expression for the optimization problem

$$F_{K,\mathbf{h},\boldsymbol{\lambda}}^* = \sup_{\mathbf{x} \in S^{K-1}} \{F_{K,\mathbf{h},\boldsymbol{\lambda}}(\mathbf{x})\} = \max_{\mathbf{x} \in S^{K-1}} \{F_{K,\mathbf{h},\boldsymbol{\lambda}}(\mathbf{x})\}. \quad (2.2)$$

**Proposition 2.1.** *For any  $\mathbf{h} \in \sqrt{\Gamma}S^{K-1}$  and  $\boldsymbol{\lambda} \in \mathbb{R}_{\geq}^K$  let  $\nu_{\mathbf{h}} = \Gamma^{-1} \sum_{i=1}^K h_i^2 \delta_{\lambda_i}$ . Then,*

$$F_{K,\mathbf{h},\boldsymbol{\lambda}}^* = F(\lambda_1, \nu_{\mathbf{h}}; \Gamma). \quad (2.3)$$

Before proving Proposition 2.1, we treat the following easier case.

**Lemma 2.2.** *Assume  $h_1 \neq 0$ . Let  $\theta^*$  be the unique solution in  $(\lambda_1, \infty)$ , of*

$$\sum_{i=1}^K \frac{h_i^2}{(\theta^* - \lambda_i)^2} = \Gamma. \quad (2.4)$$

*Then*

$$F_{K,\mathbf{h},\boldsymbol{\lambda}}^* = \frac{1}{2} \left( \theta^* + \sum_{i=1}^K \frac{h_i^2}{\theta^* - \lambda_i} \right) = \frac{1}{2} \inf_{\theta > \lambda_1} \left( \theta + \sum_{i=1}^K \frac{h_i^2}{\theta - \lambda_i} \right). \quad (2.5)$$

*Proof of Lemma 2.2.* Note first that  $F_{K,\mathbf{h},\boldsymbol{\lambda}}^* = F_{K,|\mathbf{h}|,\boldsymbol{\lambda}}^*$ , where  $|\mathbf{h}|_i = |h_i|$ . We thus may assume that  $h_i \geq 0$  for all  $i$ . By adding a constant to all  $\lambda_i$ , we may and will also assume that  $\lambda_K > 0$ . Finally, with  $\mathcal{B} := \{\mathbf{x} : \sum_{i=1}^K x_i^2 \leq 1, x_i \geq 0\}$ , one has from the monotonicity of  $a \mapsto F_{K,\mathbf{h},\boldsymbol{\lambda}}(a\mathbf{x})$  in  $\mathcal{B}$  that

$$\max_{\mathbf{x} \in \mathcal{B}} \{F_{K,\mathbf{h},\boldsymbol{\lambda}}(\mathbf{x})\} = \sup_{\mathbf{x} \in S^{K-1}} \{F_{K,\mathbf{h},\boldsymbol{\lambda}}(\mathbf{x})\}.$$

Note also that the maximum of the strictly convex continuous function  $F_{K,\mathbf{h},\boldsymbol{\lambda}}(\cdot)$  on the convex domain  $\mathcal{B}$  is obtained at a unique  $\mathbf{x}^* \in S^{K-1}$  due to the compactness of  $S^{K-1}$  and the monotonicity of  $a \mapsto F_{K,\mathbf{h},\boldsymbol{\lambda}}(a\mathbf{x})$  in  $\mathcal{B}$ .

Using the Lagrange multiplier  $\frac{\theta}{2}(\sum_{i=1}^K x_i^2 - 1)$ , we obtain that  $x_i^*(\lambda_i - \theta^*) + h_i = 0$  for all  $i$ . This gives  $x_i^* = h_i/(\theta^* - \lambda_i)$  for some  $\theta^*$  that must satisfy (2.4). Since  $x_1^* \geq 0$  is finite and  $h_1 > 0$ , this means in particular that  $\theta^* > \lambda_1$ . The monotonicity of  $\theta \mapsto \sum_{i=1}^K h_i^2/(\theta - \lambda_i)^2 =: f(\theta)$  on  $[\lambda_1, \infty)$  together with  $f(\lambda_1) = \infty$ ,  $f(\infty) = 0$  yields the uniqueness of such  $\theta^*$  satisfying (2.4), as well as the left equality in (2.5). The second part of (2.5) then follows by carrying out the optimization over  $\theta$  in the right hand side and noting that its solution  $\bar{\theta}$  must also satisfy (2.4), hence coincide with  $\theta^*$ . ■

*Proof of Proposition 2.1.* The right side of (2.5) is precisely  $F(\lambda_1, \nu_{\mathbf{h}}; \Gamma)$ . Hence, in view of Lemma 2.2, it suffices to consider the case of  $h_1 = 0$ , which we handle by approximation. That is, we set  $\mathbf{h}^\epsilon \in \sqrt{\Gamma}S^{K-1}$  so that  $h_1^\epsilon = \sqrt{\Gamma}\epsilon > 0$  and  $h_i^\epsilon = \sqrt{1 - \epsilon^2}h_i$  for all  $i \geq 2$ . Setting  $\phi(\epsilon) = \epsilon + (1 - \sqrt{1 - \epsilon^2})$ , note that

$$|F_{K,\mathbf{h}^\epsilon,\boldsymbol{\lambda}}^* - F_{K,\mathbf{h},\boldsymbol{\lambda}}^*| \leq \sqrt{\Gamma}\phi(\epsilon) \rightarrow_{\epsilon \rightarrow 0} 0.$$

Further,  $\nu_{\mathbf{h}^\epsilon} \rightarrow \nu_{\mathbf{h}}$  in  $\mathcal{M}_1$  hence  $F(\lambda_1, \nu_{\mathbf{h}^\epsilon}; \Gamma) \rightarrow F(\lambda_1, \nu_{\mathbf{h}}; \Gamma)$  as  $\epsilon \rightarrow 0$  (see part **I** of proof of Proposition 1.5), and the right side of (2.5) yields (2.3).  $\blacksquare$

**2.3. An auxiliary LDP for squares of normal variables.** We consider here an auxiliary LDP. Specifically, fixing integer  $K \geq 1$ , partition  $\{1, \dots, n\}$  to non-empty, disjoint subsets  $\mathcal{I}_n(i)$ ,  $i = 1, \dots, K$ , such that  $n^{-1}|\mathcal{I}_n(i)| \rightarrow_{n \rightarrow \infty} \mu_K(i)$ , for  $i = 1, \dots, K$ , and some probability measure  $\mu_K$  on  $\{1, \dots, K\}$ . With  $\{G_j\}_{j=1}^n$  i.i.d. standard normal random variables, define the random vectors  $\mathbf{X} = \{X_1, \dots, X_K\}$  and  $\bar{\mathbf{X}} = \{\bar{X}_1, \dots, \bar{X}_K\}$ , such that  $X_i = n^{-1} \sum_{j \in \mathcal{I}_n(i)} G_j^2$  and  $\bar{X}_i = X_i / \mathbf{X}^S$  for  $\mathbf{X}^S = \sum_{i=1}^K X_i$ . Note that the laws of  $\mathbf{X}$  and  $\bar{\mathbf{X}}$  depend on  $n$ . Finally, let  $\mathbb{S}_K = \{\mathbf{x} \in \mathbb{R}_+^K : \sum_i x_i = 1\}$  and associate to each point  $\mathbf{x} \in \mathbb{S}_K$  the probability measure  $\mu_{\mathbf{x}}$  on  $\{1, \dots, K\}$  such that  $\mu_{\mathbf{x}}(i) = x_i$ ,  $i = 1, \dots, K$ .

**Proposition 2.3.** *The random vectors  $\bar{\mathbf{X}}$  satisfy (as  $n \rightarrow \infty$ ) the LDP in  $\mathbb{S}_K$  with speed  $n$  and GRF*

$$J(\mathbf{x}) = \frac{1}{2} H(\mu_K | \mu_{\mathbf{x}}) \quad (2.6)$$

where  $H(\mu_K | \mu_{\mathbf{x}}) = \sum_{i=1}^K \mu_K(i) \log(\mu_K(i) / \mu_{\mathbf{x}}(i))$ , and we adopt the convention  $0 \log(0/x) = 0$  for all  $x \geq 0$ .

To prove Proposition 2.3, we first establish an elementary result concerning large deviations of  $\chi$ -square variables.

**Lemma 2.4.** *Suppose integers  $\ell_n \geq 1$  are such that  $n^{-1}\ell_n \rightarrow \alpha \in [0, 1]$ . Then,  $Y_n = n^{-1} \sum_{j=1}^{\ell_n} G_j^2$  satisfies the large deviations on  $[0, \infty)$  with GRF  $J_\alpha(y) = \frac{1}{2}(y - \alpha + \alpha \log(\alpha/y))$  (where again by convention  $0 \log(0/x) = 0$ ).*

*Proof.* A direct computation shows that

$$\frac{1}{n} \log E(e^{\theta n Y_n}) = -\frac{\ell_n}{2n} \log(1 - 2\theta)_+.$$

In case  $n^{-1}\ell_n \rightarrow \alpha > 0$  an application of the Gartner-Ellis theorem (see [DZ98, Theorem 2.3.6] for this version), yields the claim. On the other hand, if  $n^{-1}\ell_n \rightarrow 0$ , fix  $y > 0$  and  $\theta_n \uparrow 1/2$  slow enough for  $n^{-1}\ell_n \log(1 - 2\theta_n) \rightarrow 0$ . Then,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P(Y_n \geq y) \leq -\lim_{n \rightarrow \infty} (\theta_n y + \frac{\ell_n}{2n} \log(1 - 2\theta_n)) = -\frac{y}{2} = -J_0(y),$$

while since  $\ell_n \geq 1$ ,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log P(Y_n \geq y) \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log P(G_1^2 > ny) = -\frac{y}{2},$$

which completes the proof.  $\blacksquare$

*Proof of Proposition 2.3.* Let  $f : \mathbb{R}_+^K \mapsto [0, \infty)$  and  $g : \mathbb{R}_+^K \setminus \{0\} \mapsto \mathbb{S}_K$  be defined by  $f(\mathbf{x}) = \sum_{i=1}^K x_i$  and  $g(\mathbf{x}) = \frac{1}{f(\mathbf{x})} \mathbf{x}$ . By Lemma 2.4, and using the independence of its components, the vector  $\mathbf{X}$  satisfies in  $\mathbb{R}_+^K$  the LDP with speed  $n$  and GRF

$$\bar{J}(\mathbf{x}) := \frac{1}{2} (f(\mathbf{x}) - 1 - \log f(\mathbf{x}) + H(\mu_K | \mu_{g(\mathbf{x})})) .$$

Note that  $g(\mathbf{X}) = \bar{\mathbf{X}}$  and that for any  $\delta > 0$ , the function  $g(\cdot)$  is continuous on  $f^{-1}((\delta, \infty))$ . Since  $\lim_{\delta \rightarrow 0} \inf_{\{\mathbf{x}: f(\mathbf{x}) \leq \delta\}} \bar{J}(\mathbf{x}) = \infty$ , we conclude (from the contraction principle, see [DZ98, Theorem 4.2.1]), that  $\bar{\mathbf{X}}$  satisfies the LDP in  $\mathbb{S}_K$  with GRF  $J(\bar{\mathbf{x}}) = \inf_{\{\mathbf{x} \in \mathbb{R}_+^K: g(\mathbf{x}) = \bar{\mathbf{x}}\}} \bar{J}(\mathbf{x})$ . Clearly, such  $J(\bar{\mathbf{x}})$  is given by (2.6), completing the proof.  $\blacksquare$

**2.4. LDP for quadratic optimization - the diagonal case.** We modify the optimization problem  $F_{n, \mathbf{h}, \boldsymbol{\lambda}}^*$  so that the LDP of Proposition 2.3 can be applied. To this end, for  $k = 1, 2, \dots$ , we let  $K = K(k) = 2^k$  and form refined partitions of the intervals  $[\lambda_-^* - \delta_k, \lambda_+^* + \delta_k]$  to disjoint sub-intervals  $I_i^{(k)} = [\lambda_i^-, \lambda_i^+)$ ,  $i = 1, \dots, K$ , such that  $\lambda_i^- < \lambda_+^*$ ,  $\lambda_i^+ > \lambda_-^*$ , and  $q(\{\lambda_i^\pm\}) = 0$  for  $i = 1, \dots, K$ , while  $\Delta_k := \max_{i=1}^K (\lambda_i^+ - \lambda_i^-) \rightarrow 0$  as  $k \rightarrow \infty$  (and with  $\Delta_k \geq \delta_k$ , also  $\delta_k \rightarrow 0$ ). Let  $\mathcal{I}_n^{(k)}(i) = \{j : \lambda_j(n) \in I_i^{(k)}\}$ ,  $i = 1, \dots, K$  and for any  $\mathbf{x} \in S^{n-1}$  set  $\bar{\mathbf{x}} \in S^{K-1}$  such that  $\bar{x}_i \geq 0$  and

$$\bar{x}_i^2 = \sum_{j \in \mathcal{I}_n^{(k)}(i)} x_j^2.$$

We similarly set  $\bar{\mathbf{h}} \in \sqrt{\Gamma} S^{K-1}$  such that  $\bar{h}_i \geq 0$  and  $\bar{h}_i^2 = \sum_{j \in \mathcal{I}_n^{(k)}(i)} h_j^2$ , enforcing  $\bar{x}_i = \bar{h}_i = 0$  in case the set  $\mathcal{I}_n^{(k)}(i)$  is empty. Next, subject to the latter restriction, define

$$F_{K, \bar{\mathbf{h}}, \boldsymbol{\lambda}^\pm}^* := \sup_{\bar{\mathbf{x}} \in S^{K-1}, \bar{x}_i \geq 0} \left( \frac{1}{2} \sum_{i=1}^K \lambda_i^\pm \bar{x}_i^2 + \sum_{i=1}^K \bar{h}_i \bar{x}_i \right),$$

about the LDP of which we have the following result (whose proof is deferred to the end of this sub-section).

**Proposition 2.5.** *Fix  $k$  and non-random  $\Gamma > 0$ , taking  $\mathbf{h}$  Haar distributed on  $\sqrt{\Gamma} S^{n-1}$ , independently of  $\boldsymbol{\lambda}$ .*

(a). *The sequence  $\{F_{K, \bar{\mathbf{h}}, \boldsymbol{\lambda}^\pm}^*\}$  satisfies the LDP with speed  $n$  and GRF*

$$I_q^{H,k}(m; \Gamma) = \inf \left\{ \frac{1}{2} H(q_k | \nu_k) : m = F(\lambda_1^-, \nu_k; \Gamma) \right\}, \quad (2.7)$$

where  $q_k = \sum_{i=1}^K q(I_i^{(k)}) \delta_{\lambda_i^-}$  and  $\nu_k = \sum_{i=1}^K \nu(I_i^{(k)}) \delta_{\lambda_i^-}$  for some  $\nu \in M_1(\cup_i I_i^{(k)})$ .

(b). *For any  $m \in \mathbb{R}$ ,*

$$I_q^H(m; \lambda_\pm^*, \Gamma) = \sup_{\delta > 0} \liminf_{k \rightarrow \infty} \inf_{|m' - m| < \delta} I_q^{H,k}(m'; \Gamma). \quad (2.8)$$

*Proof of Theorem 1.3.*

(a). Note that  $0 \leq F_{K, \bar{\mathbf{h}}, \boldsymbol{\lambda}^+}^* - F_{K, \bar{\mathbf{h}}, \boldsymbol{\lambda}^-}^* \leq \frac{1}{2} \Delta_k$ . Further, by Cauchy-Schwarz

$$F_{K, \bar{\mathbf{h}}, \boldsymbol{\lambda}^-}^* \leq F_{n, \mathbf{h}, \boldsymbol{\lambda}}^* \leq F_{K, \bar{\mathbf{h}}, \boldsymbol{\lambda}^+}^*, \quad (2.9)$$

as soon as  $\lambda_n(n) \geq \lambda_-^* - \delta_k$  and  $\lambda_1(n) \leq \lambda_+^* + \delta_k$ . By Assumption 1.2, the inequality (2.9) holds for all  $n$  large enough, hence the collection  $\{F_{K, \bar{\mathbf{h}}, \boldsymbol{\lambda}^\pm}^*\}$  is an exponentially good approximation of  $\{F_{n, \mathbf{h}, \boldsymbol{\lambda}}^*\}$  (see [DZ98, Definition 4.2.14]). In view of [DZ98, Theorem 4.2.16, part (a)] (see also [DZ98, Exercise 4.2.29, part (a)]), part (a) of Theorem 1.3 is thus a direct consequence of Proposition 2.5.

(b). We represent the centered multivariate normal random vector  $\mathbf{g}$  of covariance matrix  $\frac{\Gamma}{n}\mathbf{I}_n$  as the product of Haar distributed  $\mathbf{h} \in \sqrt{\Gamma}S^{n-1}$  and the independent  $\sqrt{Y_n}$ , where  $nY_n$  has  $\chi$ -square law of  $n$  degrees of freedom. Hence,  $F_{n,\mathbf{g},\boldsymbol{\lambda}}^* = F_{n,\sqrt{Y_n}\mathbf{h},\boldsymbol{\lambda}}^*$  for  $Y_n$  of Lemma 2.4 (with  $\ell_n = n$ , so  $\alpha = 1$ ), which is further independent of  $\mathbf{h}$  and  $\boldsymbol{\lambda}$ . In particular, the exponentially tight  $\{Y_n\}$  satisfies the LDP in  $\mathbb{R}_+$  with the GRF  $J_1(y)$  of Lemma 2.4. Moreover, from (1.5) we have that  $\sqrt{y} \mapsto F_{n,\sqrt{y}\mathbf{h},\boldsymbol{\lambda}}^*$  is globally Lipschitz continuous, uniformly in  $n$ ,  $\boldsymbol{\lambda}$  and  $\mathbf{h} \in \sqrt{\Gamma}S^{n-1}$ , so upon a suitable discretization of the range of  $\sqrt{Y_n}$ , we get part (b) of Theorem 1.3 as an immediate consequence of part (a) of this theorem (for a similar argument, see [DZ98, Exercise 4.2.7]).  $\blacksquare$

*Proof of Proposition 2.5.*

(a). Fixing  $k$  and  $\Gamma > 0$ , we apply Proposition 2.1, to find that for each  $\bar{\mathbf{h}} \in \sqrt{\Gamma}S^{K-1}$ ,

$$F_{K,\bar{\mathbf{h}},\boldsymbol{\lambda}^-}^* = F(\lambda_1^-, \nu_{\bar{\mathbf{h}}}; \Gamma)$$

where  $\nu_{\bar{\mathbf{h}}} = \Gamma^{-1} \sum_{i=1}^K \bar{h}_i^2 \delta_{\lambda_i^-}$ . Next, let  $\mathcal{J}_* := \{1, K\} \cup \{1 < i < K : q(I_i^{(k)}) > 0\}$  and note that  $\mathcal{I}_n^{(k)}(i)$  is non-empty for all  $i \in \mathcal{J}_*$  and  $n \geq n_0(k)$ . Indeed, with  $\lambda_-^* < \lambda_K^+$  and  $\lambda_1^- < \lambda_+^*$ , we have from (A2) and (A3) that both  $\mathcal{I}_n^{(k)}(1)$  and  $\mathcal{I}_n^{(k)}(K)$  are non-empty for all  $n \geq n_0(k)$ , whereas by (A1) and our condition that  $q(\{\lambda_i^\pm\}) = 0$ , the same applies whenever  $q(I_i^{(k)}) > 0$ . Thus, dividing the positive integers to at most  $2^{K-2}$  possibilities, we have upon passing to the relevant sub-sequence, that for some fixed  $\mathcal{J}_* \subseteq \mathcal{J} \subseteq \{1, \dots, K\}$  and all  $n$ ,

$$n^{-1}|\mathcal{I}_n^{(k)}(i)| = L_n^\lambda(I_i^{(k)}) > 0 \iff i \in \mathcal{J}.$$

Taking  $\mathbf{h} \in \sqrt{\Gamma}S^{n-1}$  according to Haar measure, and setting  $K' = |\mathcal{J}|$ , we have that along such sub-sequence  $\Gamma^{-1}(\bar{h}_i^2, i \in \mathcal{J}) \in \mathbb{S}_{K'}$  has the law of  $\bar{\mathbf{X}}$  of Proposition 2.3, with  $\mu_K(i) = \lim_{n \rightarrow \infty} L_n^\lambda(I_i^{(k)})$  given by  $q(I_i^{(k)}) = q_k(\{\lambda_i^-\})$  (by Assumption (A1) and having  $q(\partial I_i^{(k)}) = 0$  for all  $i$ ). Now, for any fixed  $\Gamma$  and  $\{\lambda_i^-, i = 1, \dots, K\}$ , the function  $F_{K,\bar{\mathbf{h}},\boldsymbol{\lambda}^-}^*$  of  $(\bar{h}_i^2, i \in \mathcal{J})$  is continuous. Thus, along such subsequence we get the LDP in part (a) of Proposition 2.5 from Proposition 2.3 (together with the contraction principle), albeit having to take in the formula (2.7) of its GRF only  $\nu_k$  supported on  $\cup_{i \in \mathcal{J}} I_i^{(k)}$ . Further, rewriting the proof of Proposition 1.10 part (a) for  $q_k$  and  $I_q^{H,k}$  (instead of  $q$  and  $I_q^H$ ), we deduce that the GRF of (2.7) is unchanged by reducing the support  $\nu_k$ , as long as it contains  $\cup_{i \in \mathcal{J}_*} I_i^{(k)}$  (see (2.17)). This is the case here, regardless of the sub-sequence we follow, thereby completing the proof of part (a).

(b). Fixing  $\Gamma, m$  and turning to the proof of (2.8), note that every  $\nu \in M_1([\lambda_-^*, \lambda_+^*])$  of  $H(q|\nu)$  finite, induces the sequence  $\nu_k = \sum_{i=1}^K \nu(I_i^{(k)}) \delta_{\lambda_i^-}$  such that  $H(q_k|\nu_k) \uparrow H(q|\nu)$  (for example, use  $L_1(\nu)$ -approximations of the relevant bounded continuous test function  $\phi$  in the variational representation of [DZ98, Lemma 6.2.13], by simple functions based on the refined partitions  $\{I_i^{(k)}\}$ ). Further, from (1.7) it is easy to see that for any  $\xi \geq \lambda_+^*$ ,

$$F(\xi, \nu; \Gamma) \geq F(\xi, \nu_k; \Gamma) \geq F(\xi + \Delta_k, \nu; \Gamma) - \frac{1}{2} \Delta_k \quad (2.10)$$

(with the right-inequality holding as soon as  $\xi \geq \lambda_1^-$ ). Now, if  $I_q^H(m; \lambda_\pm^*, \Gamma) < \infty$ , then for any  $\epsilon > 0$  there exists  $\nu = \nu^{(\epsilon)} \in M_1([\lambda_-^*, \lambda_+^*])$  such that  $\frac{1}{2}H(q|\nu) \leq I_q^H(m; \lambda_\pm^*, \Gamma) + \epsilon$  and

$F(\lambda_+^*, \nu; \Gamma) = m$ . Setting  $m_k = F(\lambda_1^-, \nu_k; \Gamma)$  for  $\nu = \nu^{(\epsilon)}$ , the latter property yields, upon considering the left-inequality of (2.10) at  $\xi = \lambda_+^* \in [\lambda_1^-, \lambda_1^- + \Delta_k]$  and its right-inequality for  $\xi = \lambda_1^-$ , that

$$m \geq F(\lambda_+^*, \nu_k; \Gamma) \geq m_k \geq m - \frac{1}{2}\Delta_k. \quad (2.11)$$

With  $\Delta_k \rightarrow 0$ , by (2.7) and our choice of  $\nu = \nu^{(\epsilon)}$ , this implies that for some  $m_k \rightarrow m$ ,

$$I_q^H(m; \lambda_\pm^*, \Gamma) + \epsilon \geq \frac{1}{2}H(q_k | \nu_k) \geq I_q^{H,k}(m_k; \Gamma).$$

Taking now  $\epsilon \downarrow 0$ , we conclude that the l.h.s. of (2.8) exceeds its r.h.s.

For the converse direction, note that  $H(q|\nu) = H(q_k|\nu_k)$  for any given  $k$  and  $\nu_k(\cdot)$  considered in (2.7), provided  $\nu \in M_1([\lambda_-^*, \lambda_+^*])$  is given by

$$\nu(\cdot) := \sum_{i=1}^K q(\cdot | I_i^{(k)}) \nu_k(\{\lambda_i^-\}).$$

Further, (2.10) holds for this choice of  $\nu(\cdot)$ , resulting as in the derivation of (2.11) with

$$|F(\lambda_+^*, \nu; \Gamma) - F(\lambda_1^-, \nu_k; \Gamma)| \leq \frac{1}{2}\Delta_k \rightarrow 0.$$

Since this applies for any  $\nu_k$  which is considered in determining  $I_q^{H,k}(m'; \Gamma)$ , it follows that the r.h.s. of (2.8) exceeds

$$\sup_{\delta > 0} \inf_{|m' - m| < 2\delta} \{I_q^H(m'; \lambda_\pm^*, \Gamma)\} = I_q^H(m; \lambda_\pm^*, \Gamma)$$

(due to the lower semi-continuity of  $I_q^H(\cdot; \lambda_\pm^*, \Gamma)$ , which was proved in Proposition 1.5).  $\blacksquare$

**Remark 2.6.** Denoting by  $d_{\text{BL}}(\cdot, \cdot)$  the bounded-Lipschitz metric compatible with weak convergence in  $M_1(\mathbb{R})$ , let  $B_n((q, \psi^*), \eta)$  denote the collection of  $\lambda \in \mathbb{R}_{\geq}^n$  such that  $d_{\text{BL}}(L_n^\lambda, q) < \eta$  and  $|\lambda_1(n) - \psi^*| < \eta$ . In conjunction with Remark 1.6, our proof of Theorem 1.3 actually gives for any  $m \in \mathbb{R}$  the stronger, uniform conclusion in part (b),

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \liminf_{\eta \downarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \inf_{\lambda \in B_n((q, \psi^*), \eta)} P(|F_{n, \mathbf{g}, \lambda}^* - m| < \epsilon) = -I_q^G(m; \psi^*, \Gamma) \\ & = \lim_{\epsilon \downarrow 0} \limsup_{\eta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \sup_{\lambda \in B_n((q, \psi^*), \eta)} P(|F_{n, \mathbf{g}, \lambda}^* - m| < \epsilon), \quad \forall \psi^* \geq q_+. \end{aligned} \quad (2.12)$$

The same conclusion applies for the LDP for Haar distributed  $\mathbf{h}$ , of GRF  $I_q^H(m; \lambda_\pm^*, \Gamma)$  which we proved in part (a) of Theorem 1.3, except for replacing in this case  $\lambda_1(n)$  by  $\lambda_n(n)$  whenever  $m < \bar{m}$ , and considering then  $\psi^* \leq q_-$ .

## 2.5. LDP for matrices: proof of Cor. 1.7 and 1.9.

*Proof of Corollary 1.7.* Fix a sequence of symmetric  $\mathbb{R}$ -valued matrices  $\{W_n\} \in \mathcal{W}_{\lambda_\pm^*, q}$ . For each  $n$ , the matrix  $W_n$  of eigenvalue vector  $\lambda$ , is of the form  $W_n = O_n^T D_n O_n$ , for  $D_n = \text{diag}(\lambda_1, \dots, \lambda_n)$  and some real, orthogonal matrix  $O_n$ . Any such  $O_n$  induces the



isomorphism  $\mathbf{y} = O_n \mathbf{x}$  on  $S^{n-1}$ , such that  $\mathbf{h} = O_n \tilde{\mathbf{h}}$  is Haar distributed on  $\sqrt{\Gamma} S^{n-1}$ , independently of  $\boldsymbol{\lambda}$ . Further, in view of (1.2) and (1.5),

$$F_n^{W, \tilde{\mathbf{h}}} = \sup_{\mathbf{y} \in S^{n-1}} \left( \frac{1}{2} \langle \boldsymbol{\lambda}, \mathbf{y}^2 \rangle + \langle \mathbf{h}, \mathbf{y} \rangle \right) = F_{n, \mathbf{h}, \boldsymbol{\lambda}}^*.$$

Part (a) is thus an immediate consequence of part (a) of Theorem 1.3 and the definition of  $\mathcal{W}_{\lambda_{\pm}^*, q}$ . Similarly, considering the multivariate normal  $\tilde{\mathbf{g}}$  of covariance  $\frac{\Gamma}{n} \mathbf{I}_n$ , results with  $\mathbf{g} = O_n \tilde{\mathbf{g}}$  having the same law as  $\tilde{\mathbf{g}}$ , independently of  $\boldsymbol{\lambda}$ , and consequently the LDP of part (b) for  $F_n^{W, \tilde{\mathbf{g}}}$  follows from part (b) of Theorem 1.3 about the LDP of  $F_{n, \mathbf{g}, \boldsymbol{\lambda}}^*$ .  $\blacksquare$

*Proof of Corollary 1.9.* We first convert  $F_n^{W, \tilde{\mathbf{g}}}$  of law  $\mathbb{P}_{\Gamma}^{G, n}$  into  $F_{n, \mathbf{g}, \boldsymbol{\lambda}}^*$  as in the proof of Corollary 1.7, just now for random  $\boldsymbol{\lambda} \in \mathbb{R}_{\geq}^n$  having the joint eigenvalue density of the GOE. Recall that the convergence of  $L_n^{\boldsymbol{\lambda}}$  to  $\sigma$ , in  $M_1(\mathbb{R})$ , occurs with exponential speed  $n^2$  (see [BG97]). Hence, fluctuations from this convergence can not affect the LDP considered here, which is at exponential speed  $n$ . Specifically, even when proving the LDP upper bound, we can assume w.l.o.g. that  $d_{\text{BL}}(L_n^{\boldsymbol{\lambda}}, \sigma) < \eta$  for any  $\eta > 0$  and all  $n \geq n_0(\eta)$ . Further,  $\frac{1}{2} \lambda_1(n) \leq F_{n, \mathbf{g}, \boldsymbol{\lambda}}^* \leq \frac{1}{2} \lambda_1(n) + \|\mathbf{g}\|$ , where both  $\{\lambda_1(n)\}$  and  $\{\|\mathbf{g}\|\}$  are exponentially tight (due to their LDP having a GRF, see [BDG01, Theorem 6.2] and Lemma 2.4, respectively). Hence, the sequence  $\{(\lambda_1(n), F_{n, \mathbf{g}, \boldsymbol{\lambda}}^*)\}$  is exponentially tight in  $\mathbb{R}^2$ , and to establish part (b) of the corollary, it suffices to show that for any  $m \geq 1$ ,  $\psi^* \geq 2$ ,

$$\begin{aligned} & \lim_{\varepsilon, \eta \downarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log P(\boldsymbol{\lambda} \in B_n((\sigma, \psi^*), \eta), |F_{n, \mathbf{g}, \boldsymbol{\lambda}}^* - m| < \varepsilon) = -[I_{\sigma}^G(m; \psi^*, \Gamma) + I_e(\psi^*)] \\ & = \lim_{\varepsilon, \eta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P(\boldsymbol{\lambda} \in B_n((\sigma, \psi^*), \eta), |F_{n, \mathbf{g}, \boldsymbol{\lambda}}^* - m| < \varepsilon) \end{aligned} \quad (2.13)$$

(this is enough due to general considerations, c.f. [DZ98, Theorems 4.1.11 and 4.2.1]). To this end, fix  $m \geq 1$  and  $\psi^* \geq 2$ . Since the events considered in (2.13) are monotone in both  $\varepsilon$  and  $\eta$ , we can and will take  $\eta \downarrow 0$  before considering  $\varepsilon \downarrow 0$ . Then, writing

$$\begin{aligned} & \frac{1}{n} \log P(\boldsymbol{\lambda} \in B_n((\sigma, \psi^*), \eta), |F_{n, \mathbf{g}, \boldsymbol{\lambda}}^* - m| < \varepsilon) \\ & = \frac{1}{n} \log P(|F_{n, \mathbf{g}, \boldsymbol{\lambda}}^* - m| < \varepsilon \mid \boldsymbol{\lambda} \in B_n((\sigma, \psi^*), \eta)) + \frac{1}{n} \log P(\boldsymbol{\lambda} \in B_n((\sigma, \psi^*), \eta)), \end{aligned} \quad (2.14)$$

we have from the uniform bounds of (2.12), that the term involving the conditional probability converges to  $-I_{\sigma}^G(m; \psi^*, \Gamma)$  when  $n \rightarrow \infty$  followed by  $\eta \downarrow 0$  and finally  $\varepsilon \downarrow 0$ . Further, due to the much stronger concentration of  $L_n^{\boldsymbol{\lambda}}$  under the GOE law, for an LDP at exponential speed  $n$ , the events  $\{\boldsymbol{\lambda} \in B_n((\sigma, \psi^*), \eta)\}$  are then equivalent to  $\{|\lambda_1(n) - \psi^*| < \eta\}$ . Hence, in the limit  $n \rightarrow \infty$  followed by  $\eta \downarrow 0$ , the right-most term of (2.14) converges to  $-I_e(\psi^*)$  (by the LDP of [BDG01, Theorem 6.2] for the top eigenvalue  $\{\lambda_1(n)\}$ , under the GOE law). Combining all this, completes the proof of (2.13) and thereby of part (b) of the corollary.

Upon replacing  $I_{\sigma}^G(\cdot; \cdot)$  by  $I_{\sigma}^H(\cdot; \cdot)$ , the same argument applies in the Haar setting of  $\{\mathbb{P}_{\Gamma}^{H, n}\}_{n \geq 1}$  provided  $m \geq \bar{m}$ . However, to make use of Remark 2.6, here we must separately consider  $m < \bar{m}$ , for which the relevant rare event considered in  $B_n((\sigma, \psi^*), \eta)$  is that of having  $|\lambda_n(n) - \psi^*| < \eta$ , for fixed  $\psi^* \leq -2$  (and all  $n$  large enough). By the symmetry of the GOE law, the LDP for  $\{\lambda_n(n)\}$  is up to a sign change of its GRF, the same as the LDP for  $\{\lambda_1(n)\}$ . Adapting the preceding argument to accomodate for these additional changes,

takes care of this case as well. The GRF  $I^{H,A}(m; \Gamma)$  we thus obtain for the LDP of  $\{\mathbb{P}_\Gamma^{H,n}\}$  matches the expression (1.13), where it is optimal to set  $\psi_+^* = 2$  when  $m \leq \bar{m}$  and  $\psi_-^* = -2$  when  $m \geq \bar{m}$ .  $\blacksquare$

## 2.6. Rate functions: explicit formulas.

*Proof of Proposition 1.10.*

(a). When computing the rate function  $I_q^H(\cdot; \lambda_\pm^*, \Gamma)$  we consider only  $\nu \in M_1([\lambda_-^*, \lambda_+^*])$  such that  $q \ll \nu$ . In particular, decomposing such  $\nu = \nu_{ac} + \nu_s$  to its a.c. and singular parts with respect to  $q$ , necessarily  $\nu_{ac} = \phi q$  for some function  $\phi$  which is  $q$ -a.e. positive on  $[\lambda_-^*, \lambda_+^*]$ . Setting  $t = \nu_s([\lambda_-^*, \lambda_+^*]) = 1 - \int \phi dq$ , elementary algebra shows that

$$\frac{1}{2}H(q|\nu) = \int J_1(\phi(x))q(dx) + \frac{t}{2}, \quad (2.15)$$

$$F(\lambda_+^*, \nu; \Gamma) = \frac{1}{2} \inf_{\theta > \lambda_+^*} \left[ \theta + \Gamma \int \frac{\phi(x)}{\theta - x} q(dx) + \Gamma \int \frac{\nu_s(dx)}{\theta - x} \right], \quad (2.16)$$

with  $I_q^H(m; \lambda_\pm^*, \Gamma)$  given by minimizing the r.h.s of (2.15) over non-negative  $\phi$  and  $q$ -singular, non-negative measure  $\nu_s$  of total mass  $t = 1 - \int \phi dq$ , subject to the given value  $m$  of the r.h.s. of (2.16). The r.h.s. of (2.15) increases in  $t$ , in  $(\phi - 1)_+$  and in  $(1 - \phi)_+$ , with the global minimum (zero) attained at  $\phi = 1$  and  $t = 0$ , for which the expression (2.16) equals  $\bar{m}$ . Thus, the optimal choice is  $\nu_s = t\delta_{\psi^*}$  with  $\psi^* = \lambda_+^*$  for  $m \geq \bar{m}$  and  $\psi^* = \lambda_-^*$  for  $m \leq \bar{m}$ . That is,

$$I_q^H(m; \lambda_\pm^*, \Gamma) = \inf \left\{ \int J_1(\phi) dq + \frac{t}{2} : m = F(\lambda_+^*, \phi q + t\delta_{\psi^*}; \Gamma) \right\}, \quad (2.17)$$

where we require that  $0 \leq t = 1 - \int \phi dq$ . Adding to the r.h.s. of (2.15) the Lagrange multiplier

$$A[F(\lambda_+^*, \phi q + t\delta_{\psi^*}; \Gamma) - m] + \frac{(B-1)}{2} \left[ \int \phi(x)q(dx) + t - 1 \right], \quad (2.18)$$

we find that the infimum (over  $\phi$ ), is attained for some  $\phi^*(x) = (\theta - x)/(B\psi - Bx)$  (with the equality holding  $q$ -a.e. and  $B\psi = B\theta + A\Gamma$ ). Further, per  $\phi$  and  $t$ , the value of  $F(\lambda_+^*, \phi q + t\delta_{\psi^*}; \Gamma)$  is attained either at the unique  $\theta > \lambda_+^*$  for which

$$D(\theta) := \Gamma \int \frac{\phi(x)}{(\theta - x)^2} q(dx) + \frac{\Gamma t}{(\theta - \psi^*)^2} = 1, \quad (2.19)$$

or at  $\theta = \lambda_+^*$ , in case  $D(\lambda_+^*) \leq 1$ . Now, by our assumption that  $q_\pm = \lambda_\pm$ , the positivity of  $\phi^*(\cdot)$  requires  $\psi \geq \lambda_+^*$ ,  $B > 0$ , or  $\psi \leq \lambda_-^*$ ,  $B < 0$  (or  $B\psi = A\Gamma > 0$  when  $B = 0$ ), and with our Lagrange multiplier we find that  $t = 0$  is optimal unless  $\psi = \psi^* = \lambda_+^*$ ,  $B > 0$  or  $\psi = \psi^* = \lambda_-^*$ ,  $B < 0$ . The constraint  $t = 1 - \int \phi^* dq$  amounts to (1.19) and after some algebra we deduce that  $\phi^*(x)$  results with rate function as in (1.18), where per  $m$  (and  $B$  satisfying (1.19)), the values of  $\theta, \psi, t$  are determined out of (1.20) (the constraint involving  $m$  in (2.17), in case  $\phi = \phi^*$ ), and (1.21) (which amounts to plugging  $\phi = \phi^*$  in (2.19)). Lastly, as claimed, for  $m > \bar{m}$  we only consider  $\psi^* = \lambda_+^*$ ,  $B > 0$  and  $\psi \in [\psi^*, \theta]$ , for which  $\phi^*(x)$  is increasing on  $[\lambda_-^*, \lambda_+^*]$ , whereas  $m < \bar{m}$  requires  $\psi^* = \lambda_-^*$  with either  $B > 0$ ,  $\psi > \theta$ , or  $B < 0$ ,  $\psi \leq \psi^*$ , in both of which cases  $\phi^*(x)$  is decreasing on  $[\lambda_-^*, \lambda_+^*]$ .

(b). The only difference between  $I_q^G(m; \lambda_\pm^*, \Gamma)$  and  $I_q^H(m; \lambda_\pm^*, \Gamma)$  is that any  $\nu(\mathbb{R}) > 0$  is allowed in the former, so here  $t \geq 0$  and  $\int \phi dq \geq 0$  are no longer constrained to sum to

one. Consequently,  $I_q^G(m; \lambda_\pm^*, \Gamma)$  is also given by the r.h.s. of (2.17), just minimizing now over  $\phi \geq 1$ ,  $\psi^* = \lambda_+^*$  and  $t \geq 0$  in case  $m > \bar{m}$ , otherwise fixing  $t = 0$  and minimizing over  $\phi \in (0, 1]$ . We proceed as in part (a), except for fixing hereafter  $B = 1$  in the Lagrange multiplier of (2.18). Apart from this fixation of  $B$ , it yields the same form of  $\phi^*(x)$ , requiring  $t = 0$  unless  $\psi = \psi^*$  and having  $\theta$  determined by (2.19). Also here if  $m > \bar{m}$  then we must have  $\psi \in [\lambda_+^*, \theta)$  with  $t = 0$  whenever  $\psi > \lambda_+^*$ , while  $\psi \in (\theta, \infty)$  and  $t = 0$  when  $m < \bar{m}$ . Finally, after some algebra we deduce that  $\phi^*(x)$  results with rate given by (1.22), for  $\theta, \psi, t$  that are determined out of (1.20) and (1.21).  $\blacksquare$

*Proof of Proposition 1.11.*

(a). We are to solve the equations (1.20)–(1.22) for  $B = 1$ , some  $\theta, \psi \geq \psi^* = \lambda_+^* = 2$  and the semi-circle law  $\sigma$ . That is, when  $\mathbf{G}(\xi) = \frac{1}{2}[\xi - \sqrt{\xi^2 - 4}]$  for  $\xi \geq 2$ . Here  $\xi \mapsto 1/\mathbf{G}(\xi) = \frac{1}{2}[\xi + \sqrt{\xi^2 - 4}]$  is monotone increasing so we can and will change variables to  $\alpha := 1/\mathbf{G}(\theta) \geq 1$  and  $\beta := 1/\mathbf{G}(\psi) \geq 1$ , denoting solutions by  $(\alpha_q, \beta_q)$ . We note that  $\alpha \geq \beta$  iff  $\theta \geq \psi$ , which holds iff  $m \geq \bar{m}$ , and express all quantities appearing in the system (1.20)–(1.22) in terms of  $(\alpha, \beta)$ . To this end, since  $\xi = \mathbf{G}(\xi) + 1/\mathbf{G}(\xi)$  we have that  $\theta = \alpha + 1/\alpha$  and  $\psi = \beta + 1/\beta$ . Further, differentiating we find that  $d\xi/d\mathbf{G}(\xi) = 1 - \mathbf{G}(\xi)^{-2}$  and hence

$$\mathbf{L}(\psi) - \mathbf{L}(\theta) = \int_\theta^\psi \mathbf{G}(\xi) d\xi = \int_{1/\alpha}^{1/\beta} g(1 - g^{-2}) dg = \frac{1}{2}(\beta^{-2} - \alpha^{-2}) - \log\left(\frac{\alpha}{\beta}\right). \quad (2.20)$$

Combining this with (1.22) yields the formula  $I_\sigma^G(m; \pm 2, \Gamma) = \mathcal{I}(\alpha, \beta) + t/2$  in terms of  $\mathcal{I}(\cdot, \cdot)$  of (1.24). Turning to determine  $(\alpha_q, \beta_q)$  out of (1.20) and (1.21), we have the following three cases to consider.

**Case I.** If  $\beta_q > 1$  then  $t = 0$  and the unique solution of (1.21) is  $\alpha_q = \max(\beta_q^{-1}(1 + \Gamma), 1)$ . Substituting into (1.20) the option  $\alpha_q = 1$  results with  $\beta_q = \frac{\Gamma}{2(m-1)}$ . However, such a solution can only be relevant if

$$1 \geq \beta_q^{-1}(1 + \Gamma) = 2(m-1)(\Gamma + 1)/\Gamma,$$

i.e. for  $m \in (1, m_L]$  as in (1.27).

**Case II.** For  $\beta_q > 1$  and  $m > m_L$  we thus must have  $\beta_q = (1 + \Gamma)/\alpha_q$ , which in view of (1.20) results with  $\alpha_q > 1$  such that

$$2m - \alpha_q^{-1} - \alpha_q = \frac{\Gamma \alpha_q}{1 + \Gamma}.$$

This amounts to  $\alpha_q > 1$  that solve the quadratic equation

$$m_c^2 \alpha^2 - 2m\alpha + 1 = 0, \quad (2.21)$$

yielding the value of  $\alpha_q$  provided in (1.28). Recall our assumption that the corresponding  $\beta_q > 1$ , i.e. that  $\alpha_q < 1 + \Gamma$ , which for  $\alpha_q$  as given in (1.28) is equivalent to  $m \in (m_L, m_U)$ .

**Case III.** By now we know that for  $m \geq m_U$  the only possible solution is  $\beta = \beta_q = 1$  (i.e.  $\psi = 2$ ), for which (1.20) provides the value of  $t = \mathfrak{T}(\alpha) \geq 0$  as stated in (1.24). In this case, upon summing (1.20) and (1.21) we deduce that  $\alpha = \alpha_q$  must satisfy the equality

$$0 = 2m + \psi - 2\theta - \Gamma \mathbf{G}(\theta) = 2\left[m + 1 - \alpha - \left(1 + \frac{\Gamma}{2}\right)\alpha^{-1}\right].$$

The unique  $\alpha \geq 1$  that solves this quadratic equation is given for  $m \geq m_U$  by  $\alpha_q$  of (1.29). Collecting together **Cases I, II** and **III**, yields the stated formula of (1.26).

(b). Clearly,  $I^{G,A} = I_\sigma^G$  for all  $m \leq \bar{m}$ , since  $F_n^{W,h}$  is an increasing function of  $\lambda_1(n)$ . We claim that  $I^{G,A} = I_\sigma^G$  also for  $m > \bar{m}$ , except when setting  $B = 1$  and  $\psi^* > 2$  in (1.20)-(1.21), results with  $t > 0$ . Indeed, adding the relevant term  $I_e(\psi^*)$  to the rate function  $\int J_1(\phi)dq + t/2$  of (2.17) and using again the Lagrange multiplier (2.18) for  $B = 1$ , optimality of  $\psi^* > 2$  requires having

$$I_e'(\psi^*) + \frac{A\Gamma t}{2(\theta - \psi^*)^2} = 0, \quad (2.22)$$

which with  $I_e'(\cdot)$  strictly positive, implies having  $t > 0$ . Next, recall from the proof of part (b) of Proposition 1.10 that  $t > 0$  requires  $\psi^* = \psi = \theta + A\Gamma$ . Since  $m > \bar{m}$  we further require that  $\theta > \psi$  and thus, from (2.22) deduce that  $t = 2(\theta - \psi)I_e'(\psi)$ . Plugging such value of  $t$  into (1.20) and (1.21) yields that  $\theta > \psi > 2$  must be such that

$$2m - \theta = \Gamma[\mathbf{G}(\psi) + 2I_e'(\psi)], \quad (2.23)$$

$$\psi - \theta = \Gamma[\mathbf{G}(\theta) - \mathbf{G}(\psi) - 2I_e'(\psi)]. \quad (2.24)$$

Next, with  $2I_e'(\xi) = \sqrt{\xi^2 - 4} = 1/\mathbf{G}(\xi) - \mathbf{G}(\xi)$ , the identities (2.23)-(2.24) are in terms of  $\alpha = 1/\mathbf{G}(\theta)$  and  $\beta = 1/\mathbf{G}(\psi)$ , equivalent to

$$\begin{aligned} 2m - \alpha^{-1} - \alpha &= \Gamma\beta, \\ \beta^{-1} + \beta - \alpha^{-1} - \alpha &= \Gamma\alpha^{-1} - \Gamma\beta. \end{aligned}$$

Up to the change  $\beta \mapsto \beta^{-1}$ , these are exactly the equations which determined  $(\alpha_q, \beta_q)$  in **Case II** of part (a). In conclusion, having  $t > 0$  requires that we take for  $\alpha$  the solution  $\alpha_a > 1$  of the quadratic equation (2.21) (which is given in (1.28)), and then set  $\beta_a^{-1} = \beta_q = (1 + \Gamma)/\alpha_a$  for the value of  $\beta$ . Such solution is only possible if  $\beta_a > 1$  or equivalently  $\alpha_a > 1 + \Gamma$ . As we have seen before in **Case II** of part (a), this amounts to  $m > m_U$ . Next, similarly to the derivation of (2.20), we find that

$$I_e(\psi) = \frac{1}{2} \int_1^{1/\beta} (g^{-1} - g)(1 - g^{-2})dg = \frac{1}{4}(\beta^2 - \beta^{-2}) - \log\beta. \quad (2.25)$$

Further, plugging in (1.22) the optimal

$$t = 2(\theta - \psi)I_e'(\psi) = (\alpha + \alpha^{-1} - \beta - \beta^{-1})(\beta - \beta^{-1}), \quad (2.26)$$

yields by (2.20) and (2.24), that for  $m > m_U$

$$\begin{aligned} I_\sigma^G(m; \pm\psi^*, \Gamma) &= \frac{1}{2}[(\theta - \psi)(\mathbf{G}(\psi) + 2I_e'(\psi)) + \mathbf{L}(\psi) - \mathbf{L}(\theta)] \\ &= \frac{1}{2}\left[(\alpha + \alpha^{-1} - \beta - \beta^{-1})\beta + \frac{1}{2}(\beta^{-2} - \alpha^{-2}) - \log\left(\frac{\alpha}{\beta}\right)\right], \end{aligned} \quad (2.27)$$

in terms of  $\alpha = \alpha_a$  and  $\beta = \beta_a$ . Summing the r.h.s. of (2.25) and (2.27), leads after some algebra to the expression  $\mathcal{I}(\alpha, \beta^{-1})$ . We have just shown that at  $m > m_U$  and  $(\alpha, \beta^{-1}) = (\alpha_a, \beta_a^{-1})$  given by the r.h.s. of (1.31), this is precisely the value of  $I^{G,A}(m; \Gamma)$  (as stated in (1.30)).

The function  $m \mapsto I^{G,A}(m; \Gamma)$  is clearly smooth everywhere except at  $m = m_L$ . It is further easy to confirm that both  $I^{G,A}(m; \Gamma)$  and its first derivative are continuous at  $m = m_L$  (where the value of this function is  $\frac{1}{2}(-\log(1 - \eta) - \eta - \frac{1}{2}\eta^2)$  and its derivative equals  $-\eta$ , for  $\eta = \Gamma/(1 + \Gamma)$ ), with the second derivative of  $I^{G,A}(m; \Gamma)$  being positive everywhere, thereby verifying its strict convexity.  $\blacksquare$

## REFERENCES

- [AGZ10] G. W. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University press, Cambridge, 2010.
- [BY88] Z. D. Bai and Y. Q. Yin, Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix, *Annals Probab.* **16** (1988), pp. 1729–1741.
- [BG97] G. Ben Arous, A. Guionnet, Large deviations for Wigner’s law and Voiculescu’s non-commutative entropy. *Probab. Theory Related Fields* **108** (1997), pp. 517–542.
- [BDG01] G. Ben Arous, A. Dembo, A. Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields* **120** (2001), pp. 1–67.
- [DZ98] A. Dembo, O. Zeitouni, *Large Deviations techniques and Applications*, 2nd ed., Springer, New-York, 1998.
- [FLD13] Y. V. Fyodorov, P. Le Doussal, Topology trivialization and large deviations for the minimum in the simplest random optimization. *J. Stat. phys.* **154** (2013), pp. 466–490.